DOCUMENT RESUME

ED 144 600 IR 005 228

AUTHOR Lamb, Connie
TITLE Margedos: A Marc-Like Format for Genealogy with Provision for Documentation.
PUB DATE Dec 76
NOTE 72p.; Best copy available

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Computer Programs; Data Processing; *Family Background; Information Processing; *Machine Translation
IDENTIFIERS *Genealogy

ABSTRACT

The purpose of this project was to create a machine-readable format for genealogical data. The Library of Congress MARC (Machine Readable Cataloging) bibliographic system was used as the pattern in producing the format. Tags and codes were used to uniquely identify the components of the record for computer manipulation. The resultant structure, called MARGEDOS (Machine Readable Genealogy with Documentation of Sources), was established for use with an expanded Latter-day Saint (LDS) Family Group Record form that provides source information for each element of data. The inclusion of source documentation for every data element in a detailed and standardized manner was done to increase the credibility of genealogical records. The project was accomplished in three phases: the development of the machine-readable format, the conversion of a completed Family Group Record into that format, and the testing of the format through computer execution. An ALGOL program was constructed to recall the genealogical data with the sources and to display and print out the record in a particular arrangement. The format performed adequately for its intended purpose and the trial run on the computer was successful. Included in the document are a sample LDS Family Record form, a sample MARGEDOS Record, and the computer program. (Author/JAB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED144600

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

MARGEDOS: A MARC-LIKE FORMAT FOR GENEALOGY WITH PROVISION FOR DOCUMENTATION

BEST COPY AVAILABLE

by

Connie Lamb

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Connie Lamb

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

A Research Paper

Submitted to The

School of Library and Information Sciences

Brigham Young University

Provo, Utah

In Partial Fulfillment

of the Requirements of

L. I. S. 697

December, 1976

IR005228

ABSTRACT

The purpose of this project was to create a machine-readable format for genealogical data. The Library of Congress MARC bibliographic system was used as the pattern in producing the format. Tags and codes were used to uniquely identify the components of the record for computer manipulation. The resultant structure, called MARGEDOS, was established for use with an expanded LDS Family Group Record form that provides source information for each element of data. The inclusion of source documentation for every data element in a detailed and standardized manner was done to encourage accurate verification and to increase the credibility of genealogical records.

To test the format, a complete Family Group Record, with documentation, was converted into the machine-readable MARGEDOS structure and keyed into a computer. An ALGOL program was constructed to recall the genealogical data with sources and to display and print out the record in a particular arrangement. The format performed adequately for its intended purpose and the trial run on the computer was successful. The format is flexible and can be expanded and adapted as needed.

PREFACE

This paper was written to encourage interest in and suggest a means of increasing the accuracy of genealogical research. The chosen method of organizing the researched data is based upon MARC, a machine-readable bibliographic format already in existence and proven to be most useful. The project entails the augmentation of the MARC concept to the verification of sources used to document genealogical information. The author is aware of the widespread interest in genealogical research and hopes that the idea and results presented in this paper will be considered by genealogical organizations as a possible system to aid in genealogy work. Because the author is a member of the Church of Jesus Christ of Latter-day Saints and familiar with its programs, this project employs the church genealogical form, but the results could be adapted to other forms. The paper provides background information on genealogy and the MARC system and then presents the suggested genealogical format.

ACKNOWLEDGMENTS

Sincere appreciation is expressed to Mr. Keith H. Stirling without whose suggestion and constant help the project could not have been done. His encouragement, ideas, and patient assistance were invaluable to the accomplishment of this research project. Gratitude is also expressed to Mrs. Dorothy M. Shields, Dr. H. Thayne Johnson and students in my research project classes for their helpful suggestions.

The personnel at the LDS Church Genealogical Department and computer consultants at Brigham Young University were accommodating and provided aid during the project. Also of assistance was the use of the BYU School of Library and Information Sciences terminal and computer time.

Appreciation is given to Verena Eastley for her conscientious effort in typing the paper. Special thanks are also expressed to my parents, Mr. and Mrs. Delbert M. Lamb, for their support and help.

TABLE OF CONTENTS

Page

PREFACE ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

Chapter

1. THE PROBLEM 1

INTRODUCTION AND PROBLEM STATEMENT 1

DEFINITIONS OF TERMS 2

DISCUSSION OF THE STUDY AND ITS IMPORTANCE 4

2. THE REVIEW OF RELATED LITERATURE 7

INTRODUCTION 7

GENEALOGY 8

FAMILY GROUP RECORD 10

AUTOMATION IN GENEALOGY 11

THE MARC SYSTEM 12

3. METHODOLOGY 16

4. MARGEDOS FORMAT STRUCTURE 19

INTRODUCTION 19

LEADER 20

RECORD DIRECTORY 21

VARIABLE FIELDS 21

SOURCES 27

Chapter Page

5. IMPLEMENTATION AND DISCUSSION 30

6. CONCLUSION 35

APPENDIXES

A. FAMILY GROUP RECORD 37

B. MARC FORMAT STRUCTURE 39

C. SAMPLE MARGEDOS RECORD 49

D. TYPES OF SOURCES 51

E. COMPUTER PROGRAM 56

SOURCES CONSULTED 58

LIST OF TABLES

Table		Page
1.	Fields, Data Elements and Identifying Codes	22
2.	Classes of Sources	23
3.	Types of Sources and Specifications	52

LIST OF FIGURES

Figure		Page
1.	Printout Resulting from Computer Execution on a MARGEDOS Record	31
2.	MARC Format Structure	44
3.	Leader	45
4.	Record Directory	46
5.	Variable Fields	47
6.	Sample MARC Record	48

Chapter 1

THE PROBLEM

Introduction and Problem Statement

"The desire to know about one's family is as old as recorded history."[1] Currently, there is widespread and earnest interest in genealogy. A very important aspect of genealogical record-keeping is the accuracy of the information. Specific references to authoritative sources make the record more reliable and more beneficial to others. Therefore, it was suggested by Mr. Keith Stirling, Assistant Professor of the School of Library and Information Sciences at Brigham Young University, that a format for genealogy be devised that could be manipulated by the computer and would allow for inclusion of source references for each piece of information.

The purpose of this project is to develop a MARC-like (machine-readable) format for the current Latter-day Saint (LDS) Family Group Record which will extend the present form so that every name, place, and date can be documented. The study will include dividing the family group form into its component elements and coding the record in a MARC-like format which would allow for

---

[1] American Society of Genealogists, Genealogical Research Methods and Sources, Vol. 1, Milton Rubincam, ed. (Washington, D.C.: n.p., 1960), p. 5.

computer processing of the genealogical records. No cost analysis or comparison of systems will be attempted.

Definitions of Terms

MARC system. MARC is an acronym for MAchine Readable Cataloging. The Library of Congress developed the MARC system which enables the cataloging data of any given work to be read and manipulated by the computer. Each record is broken down into its component parts, and each part is given a symbol which the machine is programmed to "recognize" when action on it is required. For example, a work may have the main author of John Brown. If the input is entered as $aBrown, John, the computer is programmed to read the symbol, $a, to mean that a personal name follows. It will then manipulate this piece of information as required. The breakdown of each record and the tagging structure of MARC is detailed and practical, giving the computer as much flexibility as possible for handling the data.[2]

Format. The structure of a machine-readable record. It is a means whereby a computer can discriminate between data elements when they have been translated into a machine-readable pattern. The machine format, which is a method of organizing data so that the record components can be identified, is basic to MARC.[3]

---

[2]Ruth Irvine, MARC for Cataloguers: an Explanation of Its Use (Automatic Project Report No. 4, University of Southampton, 1972), p. 1.

[3]U.S., Library of Congress, MARC Development Office, Information on the MARC System (Washington: Library of Congress, 1973), p. 2.

Data element. A single unit or piece of information contained within a record.

Record. A group of related information units or data elements given in a particular framework. In this project, the records of concern are those containing genealogical data, specifically the Family Group Record.

Genealogy. The study of recorded events in the lives of individuals. Genealogical research is the method used to identify individuals from the recorded events, identify ancestors, and chart family relationships.[4] Jones, et al.[5] describe genealogy as the "systematic study of the descent of persons and families to form a continuous chain of complete, correct, and connected family units from the present back in time to the common ancestor."

Family Group Record. A genealogical form (also called a sheet) which provides space for information on two parents and details of their children. The information includes full names, birth, marriage, and death dates and places, LDS ordinance data, relationships, and general source data (see Appendix A).

---

[4]Laureen Jaussi and Gloria D. Chaston, Fundamentals of Genealogical Research (Salt Lake City: Deseret Book, 1972), pp. 1-3; see also Theodore M. Burton, "I Have a Question," Ensign, 5 (March, 1975), 32.

[5]Vincent L. Jones, Arlene H. Eakle, and Mildred H. Christensen, Genealogical Research; a Jurisdictional Approach (Rev. ed.; Salt Lake City: Publishers Press, 1972), p. 5.

Source. The point of origin or procurement; a firsthand document or primary reference work.

Original records. Those in which the first recording of a given event in that particular type of record is made. A copy or transcription of a record is not an original record.[6]

MARGEDOS. The format created during this project; an acronym for MAchine-Readable GEnealogy with DOcumentation of Sources.

Discussion of the Study and Its Importance

During the past few years, genealogy has won growing recognition and has taken on the aspects of a professional discipline. Many persons are becoming involved in genealogical research and various societies have been founded to assist them. Never before has there been so much interest shown in the work. With this increased interest and movement toward professionalism has come significant efforts to upgrade genealogical research standards.[7]

An important assignment in the LDS Church genealogy program is the four-generation program for which a person should prepare fifteen Family Group Record forms. These forms and the information contained therein are basic to further genealogical work. A major purpose of the four-generation program is to provide church members with the opportunity to verify the accuracy and completeness of

---

[6]Derek Harland, Genealogical Research Standards (Salt Lake City: Bookcraft, 1963), p. 45.

[7]Jones, op. cit., p. 3.

existing family genealogical records.[8] Unfortunately, many forms are just copies or hearsay and have not been carefully verified and documented.

It is not the responsibility of the LDS Genealogical Department to check the correctness and completeness of submitted family group records by comparison with available sources. The department evaluates the records they receive for consistency with research and record-keeping standards only. It is, then, the responsibility of the patron to follow these standards of completeness and accuracy in order to ensure the validity of the records.[9] Fudge and Smith[10] state that ". . . the society [department] does not guarantee the accuracy of the information appearing on any sheet. Responsibility for accuracy rests with the compiler." As an aid to increase the dependability of such records, it is suggested that the present group forms be expanded to include an area after each piece of information where the compiler can note the source.

Because of the benefits of computer application, it was decided that converting the Family Group Record into machine-readable form, including the complete source data, would provide a flexible

---

[8]Church of Jesus Christ of Latter-day Saints, Priesthood Genealogy Committee, A Continuing Priesthood Program for Family Exaltation [Salt Lake City] 1970, pp. 6-7. (The Church is hereafter referred to as LDS Church.)

[9]LDS Church, Genealogical Society, Ten Basic Steps in Genealogy (Salt Lake City: Deseret Book, 1965), p. 62.

[10]George H. Fudge and Frank Smith, LDS Genealogist's Hand-Book; Modern Procedures and Systems (Salt Lake City: Bookcraft, 1972), pp. 98, 100.

and serviceable instrument in genealogical research. A MARC-like system would provide a means of organizing and identifying the data so that it could be entered into a computer and rapidly accessed with a computer program. MARC is a communications format which means that it is standardized and so facilitates the transmission of data. This type of machine-readable format was chosen for the project because it is based on a flexible file structure designed to hold any type of bibliographic record and can be adapted to other kinds of data. Not only would such a procedure aid in accessibility to the information contained in the records, but the computer could also be used to quickly check the family group forms for source material.

Mr. Stirling, in a letter written to Mr. Charles Clement at the Genealogical Library in Salt Lake City, explains well the intent and purpose of this project:

> . . . developing a MARC-like record for genealogical data where the source material for each (and every) birth, baptism, marriage, death dates, etc. as well as the identification and spelling of names will be coded into the logical record. So much of the work being presently done is copied from copies. Such a record will permit a better evaluation to be made since each record carries its own bibliography.[11]

---

[11]Letter from Mr. Keith Stirling, Assistant Professor at Brigham Young University to Mr. Charles Clement, LDS Genealogical Library, April 13, 1976.

Chapter 2

THE REVIEW OF RELATED LITERATURE

Introduction

The literature search revealed no current system or plan for converting Family Group Records into machine-readable form using the MARC pattern. No plan or procedure for documenting each data element in the genealogical record was found during the literature survey.

A similar project of adapting a MARC-like format to uses other than bibliographical data was prepared by Hilda Poon at Brigham Young University in 1976. She designed a MARC-like invoicing format for serial subscribers and publishing agents. She described the purpose of her project as the creation of a format such that "data as appearing in the hundreds of different accounting invoice forms can be standardized and the communication gap between subscriber and publishing agents can be bridged."[1]

Because this current project is a new application of MARC, the literature survey consists of background information on genealogy, the Family Group Record, automation in genealogy, and the MARC system.

---

[1]Hilda Poon, "MARC-Saf, the MARC-like Invoicing Format for Subscribers and Publishing Agents" (unpublished research paper, Brigham Young University, 1976), p. 1.

Genealogy

"From the time of earliest civilizations the keeping of genealogical records or family histories has been regarded as of the greatest importance."[2] The importance of genealogical research to LDS Church members was clearly stated by Joseph Smith when he said, "The greatest responsibility in this world that God has laid upon us is to seek after our dead." (Teachings of the Prophet Joseph Smith, p. 356)

There are many records or sources from which one can obtain genealogical information or evidence. Harland[3] separates them into three major classes: 1) personal knowledge and tradition and private family records, 2) public or official records, and 3) printed sources which have been compiled and based upon information obtained from one or both of the first two classes.

During the research process, the genealogist should investigate primary or original sources. This may be oral testimony or private, public and official records which were created through normal activities of our social and economic lives.[4]

The use of all available records is very important to provide evidence of the genealogical research findings. It is difficult to set a general standard for genealogical research because methods of

---

[2]L. G. Dine, The Genealogist's Encyclopedia (New York: Weybright and Talley, 1969), p. 15.

[3]Harland, op. cit., p. 33.

[4]Norman E. Wright and David H. Pratt, Genealogical Research Essentials (Salt Lake City: Bookcraft, 1967), p. 225.

compiling pedigrees vary according to time and locality of each problem. But each genealogist should aim for a high standard in all research and carry out searches that will result in complete, correct and connected records. All available sources should be searched for information pertinent to any search problem and such information should be carefully and thoroughly evaluated and analyzed to determine its value.[5] The book, Genealogical Research by the American Society of Genealogists,[6] has a good section about rules for judging evidence. This procedure of using original sources and being careful and thorough in each search is very important to the achievement of this project's goal.

In the booklet, Records Submission Manual, the LDS Church discusses valid sources for genealogical data and states that for submitting names, the priority sources are to be searched in the following order:

1) Civil (vital) registration
2) Church records
3) Family records
4) Any other single source that provides minimum identification
5) Multiple sources
6) United States (federal) and Canadian Census schedules
7) Probate records (wills)

The manual states that the best existing sources should be used in all cases when doing research.[7]

---

[5] Harland, op. cit., pp. 20-21.

[6] American Society of Genealogists, Genealogical Research, op. cit., pp. 47-48.

[7] LDS Church, Genealogical Society, Records Submission Manual (4th ed.; Salt Lake City: Publishers Press, 1973), pp. 3-5.

Family Group Record

In 1926, the Family Group Record form was designed for the use of LDS Church members submitting names for temple work. After some changes and improvements, the form was adopted in 1942 as the form on which all names for processing, both living and dead, were submitted to the Genealogical Society. Currently the emphasis is on machine processing of names submitted on individual entry forms. However, the Family Group Record form still has value in both church and personal files.[8] The four-generation program is a permanent church assignment, and it is very important for each individual to participate in gathering this data and filling out family group forms for their own and others' use.

The Church Record Archives comprises the department's copy of all Family Group Record sheets accepted from patrons up to 1970. Under the new program of using individual entries (since 1970), information for temple work is stored in machine-readable form and on microfilm. These records are maintained in the genealogical libraries to assist others in their research and to supplement the Temple Records Index Bureau and the Church Records Archives.[9]

It is suggested by the LDS Church that the single best source be used for information on a submitted entry form following the priority list given previously. However, in the research phase, which includes the four-generation program, it is necessary to use multiple sources. All possible sources should be consulted to assure

---

[8]Fudge, op. cit., p. 26. [9]Ibid., pp. 26, 98.

accuracy. In this project, various sources can be used to document the Family Group Record, but the high priority sources suggested by the church should be considered first.

Automation in Genealogy

The automated system which is now used by the LDS Church Genealogical Department was originally known as GIANT, an acronym for Genealogical Information And Name Tabulation. The objective of this system is to collect, process, and disseminate genealogical information on all deceased persons, to be achieved in gradual phases. The first phase is designed to collect the basic information about individuals and to store that data in one large file. Later phases will be designed to retrieve the information and to correlate the individual records.[10]

The forerunner of the present system actually began in 1961 with a program of extracting identifiable names with dates and places from filmed and printed parish registers from England. The individual names and corresponding data obtained from this Controlled Extraction Program are converted to machine language and entered into an electronic computer under the Records Tabulation Program. Two hard copies result from the computer processing. One is the Computer Parish Listings which is an alphabetical printout of the parishes included in the file. The other is the Parish Printout which lists persons

---

[10]World Conference on Records and Genealogical Seminar, Salt Lake City, 1969, Papers, Vol. 2 (Salt Lake City: Genealogical Society of the Church of Jesus Christ of Latter-day Saints, 1969), Area B, Sections 13-14a and 13-14b. See also LDS Church, Genealogical Society, Records Submission Manual, op. cit., pp. 45-46 and Burton, op. cit., pp. 32-33.

(and their dates) from the parish registers for which ordinance work has already been done; this is now on microfiche also. Patrons may submit names into this program, and these names, if not duplicates, will be added into the computer and the printout. The heart of the computer system is the Genealogical Mass File which contains the basic data of name of individual, date of event (birth, marriage, death, etc.), place where event took place, and relationships to other individuals. Currently this file is accessible to patrons through the printout and supplements.[11]

At the LDS Genealogical Department in Salt Lake City, much progress is being achieved in the use of automation besides the Controlled Extraction Program. In the near future the card catalog will be converted to machine-readable form with printouts. Also the LDS Church is developing the Family Entry System. Names of persons lacking enough information for temple work to be done are entered into the computer in family form. These usually come from fragmentary or multiple sources. General approximations of time and place are given and siblings are used as identifiers. Because of the large amount of genealogical information possessed by the department, there is no current plan to enter the total data into computers for patron access.

The MARC System

The exponential growth of recorded information housed in libraries has necessitated the development of new mediums and

---

[11] Ibid.

techniques for recording, storing, and processing bibliographic data. One such method is the machine-readable format of MARC. It was developed by the Library of Congress and has been a widespread stimulus to library automation. Through this program, libraries can obtain cataloging data in machine-readable form from a central source (Library of Congress) and use these machine records to produce various products such as catalog cards, book catalogs, bibliographies, and spine labels.[12]

Basic to MARC is its machine format which organizes the data so that the parts of the record can be identified. The machine-readable record consists of a continuous string of characters, so each part must be explicitly identified for machine manipulation. Data elements are the basic unit of information, and a collection of elements is called a field. The total record is the group of fields treated as a unit. There are two types of fields, variable fields in which the lengths of the data elements cannot be predetermined and fixed fields with elements that are always expressed with the same number of characters. These record components are labelled for computer control. To facilitate exchange and transmission of bibliographic data, a standardized communications format is needed. MARC is such a communications format which means that an institution needs only a single translation program to convert a local format to or from the communications format.[13]

---

[12]U.S., Library of Congress, Information on the MARC System, op. cit., p. 1.

[13]Ibid., pp. 2-3.

The MARC format is based on three elements; 1) structure--the basic machine framework of the record, 2) content designators--the means by which elements in the record can be identified, and 3) content--the data recorded in the fields. The format structure consists of the leader, the record directory, and the variable fields. The leader in a machine record functions in a way similar to a leader on a film. Each MARC record leader is 24 characters and provides particular information about the record such as type, total length, or bibliographic level. The record directory is like the table of contents of a book because it shows what variable fields are in the record and gives their location. Each record directory is 12 characters long and contains a three digit number which identifies a variable field and other numbers indicating the total length of the field and the starting character position where the field begins in the string of record data. The variable fields are separated and tagged according to the function they serve. These fields are composed of indicators, subfield codes which designate each element in the field, the data itself, and a field terminator which indicates the end of a field. Tags, indicators, and subfield codes function as content designators. More information and diagrams about the MARC structure are given in Appendix B.

MARC is not a computer language program that will recall the data, but rather a means of organizing and identifying pieces of information so that it can be recalled by a computer program. This same arrangement may be made for genealogical data so that it can be controlled by a computer much the same as a person reads the

Family Group Record. Because MARC is flexible, it lends itself to genealogical application for records automation.

Chapter 3

METHODOLOGY

A machine-readable format for the LDS Family Group Record was created patterned after the MARC bibliographical structure. The project was accomplished in three phases: the development of the machine-readable format, the conversion of a completed Family Group Record into that format, and the testing of the format through computer execution.

The family Group Record form was broken down into its component parts. Large areas of related data were grouped together into fields and each was given an identifying numerical tag. The fields used are husband, wife, each child (with required data for each), and submission data. Because the information requested is the same for each sheet, it was not necessary to standardize the form. However, the length of each record may vary and every possible blank is not filled every time. So the format was made flexible to accomodate variations in length and allow for all possibilities of recorded data.

Within the fields are data elements that are separate but related. Each element of information was assigned a code that relates to its meaning and uniquely identifies it.

This format was based on the assumption of an expanded Family Group Form in which every element of data is documented. So in creating the format for the machine-readable record, a code was also

designated for the source of each element and which corresponds to the data element code and its meaning. The project also included the determination of how to standardize the recording of sources, what information to include, and other editing procedures and decisions.

Following the determination of the variable fields and the assigning of tags and codes to them and their elements, the overall format structure was devised. Structures of the leader and the record directory which are necessary to control and manipulate the data were formulated. The choice of tags, codes, and symbols, as well as the determination of what information to include, how to handle source data, and the total structure organization were major decisions made during the creation of the format.

A Family Group Record was filled out by checking primary sources (original records, copies of the same, or persons to whom the events had occurred). This was written out in a string of characters in the machine-readable format including all the identifying codes and exact spacing. A leader and record directory corresponding to this record were produced by counting the characters in each field.

The terminal in the Brigham Young University School of Library and Information Sciences was used to create the file into the computer from the completed Family Group Record in machine-readable form. A program in ALGOL was developed to recall the data from the computer in a predetermined arrangement which appeared on the terminal screen and was obtained in hard copy.

The resultant MARC-like format created for genealogy was named MARGEDOS (or MARGE for short) which is an acronym for

MAchine-Readable GEnealogy with DOcumentation of Sources. The name indicates that the two significant aspects emphasized in the project were the machine-readable format developed for genealogical records and the source referencing of each element of data.

Chapter 4

MARGEDOS FORMAT STRUCTURE

Introduction

The result of the current project is the MARGEDOS format. Its structure is composed of three parts. The overall structure follows the MARC pattern and contains a leader, record directory, and variable fields, but there are no control fields. The three segments of the MARGEDOS format are identified briefly as follows:

Leader. A set of characters that provides specific information about the ensuing record. It briefly gives the general make-up of the record as well as the basic address of the data in the computer.

Record Directory. A series of units, each corresponding to a field in the record. It contains the identification tag, which indicates the meaning of the respective field. Also given are the length and starting character position of each variable field in the string of data so that any field can be extracted from the record.

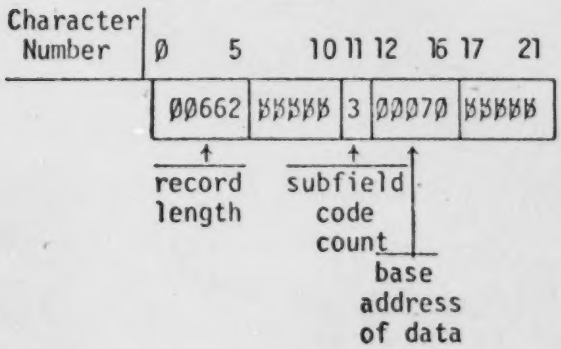
Variable Fields. A variable length record that contains the alphanumeric data from the Family Group Record and the source references. Each field is coded with a tag, as given in the record directory, which represents the meaning of the field contents. The sources for the elements in each field are listed following the data. The subfield codes and source codes, used to identify and separate

data elements, correspond to each other and link each respective element to its source.

Leader

The leader for MARGEDOS varies slightly from the MARC format because of the nature of the record, but it performs the same function. The leader is 21 characters long and is divided into several fixed length subunits. Specifically, the initial five characters indicate the length of the total record. This means that the format has the potential of incorporating a record (including data, source, and codes) up to 99,999 characters long. The next five spaces are blank providing for future expansion of the leader information. The eleventh space in the leader is the subfield code count and declares that each subfield code is three characters long. A unique subfield code precedes each data element within a field. The next five characters represent the base address which is the position into the record at which the actual data elements of the record begin. The last five spaces are again left blank for flexibility and possible addition.

Sample leader:



ß = blank

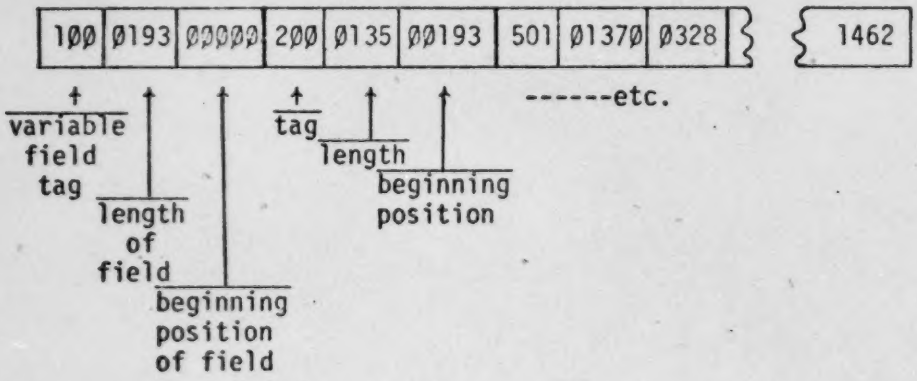
Ø = zero

Record Directory

The record directory is divided into a series of twelve character sections. Each section refers to a field and provides its

coding tag, length, and where it begins in the string of record data. The first three characters identify a particular field with a unique tag; the tags are taken from the range 001 to 999 (see Table 1). The next four digits indicate the number of characters contained in the field. Included in this count are subfield codes, the data itself, sources and their codes, and the field terminator which signifies the end of a field. The last five numeric characters give the position into the variable field of the first character associated with the corresponding tag. The first directory entry will contain all zeros in the five spaces which is the beginning position of the first variable field in the record. Subsequent record directories will have starting character positions incremented by the field length of the previous entry. All of the numbers in the record directory are right justified with zeros.

Sample record directory:



Variable Fields

In machine-readable formats, the data is entered as a string of characters, so the record is divided into fields, each identifiable

Table 1

Fields, Data Elements, and Identifying Codes

Field	Tag
Husband	100
Wife	200
Children	5XX
Submission data	800

Field 100 HUSBAND

Name of data element & code	Source code
$nh - Name of husband	@nh
$bd - Birthdate	@bd
$bp - Birthplace	@bp
$ph - Parents' names	@ph
$md - Marriage date	@md
$mp - Marriage place	@mp
$cd - Christening date	@cd
$cp - Christening place	@cp
$dd - Death date	@dd
$dp - Death place	@dp
$id - Burial date	@id
$ip - Burial place	@ip
$wa - Second wife	@wa

Table 1 (continued)

Name of data element & code	Source code
$wb - Third wife	@wb
.	.
.	.
.	.
$wnth - (--nth) wife	@wnth
$td - Baptism date	@td
$ed - Endowment date	@ed
$sd - Sealing date	@sd
$sp - Sealing place	@sp
Field 200 WIFE	
$nw - Name of wife	@nw
$bd - Birthdate	@bd
$bp - Birthplace	@bp
$pw - Parents' names	@pw
$cd - Christening date	@cd
$cp - Christening place	@cp
$dd - Death date	@dd
$dp - Death place	@dp
$id - Burial date	@id
$ip - Burial place	@ip
$ha - Second husband	@ha
$hb - Third husband	@hb
.	.
.	.
.	.
$hnth - (--nth) husband	@hnth

Table 1 (continued)

Name of data element & code	Source code
$td - Baptism date	@td
$ed - Endowment date	@ed
$sd - Sealing date	@sd
Field 5XX CHILDREN	
$nc - Name of child	@nc
$sc - Sex of child	@sc
$bd - Birthdate	@bd
$bp - Birthplace	@bp
$md - 1st Marriage date	@md
$ns - Name of spouse	@ns
$om - Other marriages (sequential order by date)	@om
$dd - Death date	@dd
$td - Baptism date	@td
$ed - Endowment date	@ed
$sd - Sealing date (to parents)	@sd
$sp - Sealing place	@sp
Field 800 SUBMISSION DATA	
$ds - Date of submission	
$sn - Submitter's name	

Table 1 (continued)

Name of data element & code
$ad - Address
$km - Stake/mission
$rh - Relationship to husband
$rw - Relationship to wife

by a unique tag given in the record directory. To simplify the MARGEDOS format and to make it more workable, only a few fields are employed. They include one field for the husband, one for the wife, one for each child, and one for submission data. The tags used are shown in Table 1. The tags for the children are noted as 5XX, and beginning with 501 for the first child, increment by one for each succeeding child. This allows for a possible 98 children and the total number is indicated by the last digit of the final 500 tag.

The data elements within each field are separated by three-character subfield codes. The format is flexible making possible the addition of further fields, data elements, and codes. The dollar sign ($) was selected as the subfield delimiter to be used with two letters; combined, they identify each element of information. The letters are generally the first letters of the data element heading such as nh for name of husband and bd for birthdate; however, a few are varied in order to avoid duplication of codes (see Table 1 for listing of codes). The same codes are used for every record and are constant across all fields, thus providing a mnemonic device to assist in creating and deciphering the machine-readable format.

The sources used to document the data elements in a particular field are listed following the data. The code used for sources is also three characters long and consists of the at sign (@) as the delimiter and two letters which are the same as those used for the corresponding data element codes (see Table 1).

The last character in each field is the field terminator. The symbol chosen as the terminator is an upward arrow (↑). See Appendix C for an example of a Family Group Record in MARGEDOS format.

Sources

The Records Submission Manual[1] lists seven types of primary sources ranked in priority sequence; these were listed previously on page 9. Restrictions regarding sources to use, when and how are also provided in the manual. This information should be consulted and followed when documenting data.

In using MARGEDOS, a standardized method of stating sources is required so as to insure proper communication of reference data. The following patterns, codes, and restrictions are employed. Primary sources are divided into nine classes based on derivation. Each class is designated by an abbreviation (see Table 2). Within each of these are several possible types of sources. Examples are given in Table 3 (see Appendix D), but the possibilities are numerous and those listed are not exhaustive. The type of source should identify specifically the record from which information was obtained, and since there are so many possibilities, it should be one or two words spelled out. The word "record" is omitted to avoid repetition. The original document should be identified even if the data is taken from a copy such as in a book or film. Two general forms are furnished for stating sources:

1) Class, type, location
2) Class, type, possessor, location

The location should provide complete place data using one of the following forms:

---

[1] LDS Church, Genealogical Society, Records Submission Manual, op. cit., pp. 3-5.

Table 2

Classes of Sources

	Class	Abbreviation
1.	Vital	VI
2.	Church	CH
3.	Civil	CV
4.	Family	FA
5.	Miscellaneous	MI
6.	Census	CE
7.	Probate	PB
8.	Personal knowledge	PK
9.	Temple	TE

1) City, county, state or province (U.S. & Canada)
2) City, state, country
3) Jurisdiction, county/state, country (e.g. parish)

There must be certain guidelines for the submitter to follow, and the information must be detailed enough to be properly coded when edited. The source data should be as complete as possible, and if original documents are not used, this should be noted. The locations and other data should be spelled out by the submitter. If needed, additional explanation to the forms stated above can be given in parentheses. The information will be abbreviated and coded during the editing process which results in a machine-readable record.

There are certain methods and specifications distinct to MARGEDOS that need to be mentioned. If a name has been changed (americanized, etc.) or the spelling altered, the alternative name may be placed in parentheses following the name used in the record. The standardized form used for dates in MARGE is day, month, year, all in arabic numerals and right justified with zeros. For temple records, "BIC" is used if "born in the covenant" is applicable. The abbreviations used in the data and sources would be determined during the editing process as the record is converted to machine-readable form. Generally they will follow post office and LDS church specifications as well as those provided in the source listing of Table 2.

Chapter 5

IMPLEMENTATION AND DISCUSSION

In addition to devising a machine-readable format, this current project included the testing of the format through computer execution. As the format was created, questions and problems of application arose resulting in revision of the format. So the tasks of developing the MARGEDOS structure and implementing it were achieved through a trial and error process. Much thought and consideration were given to the various aspects of the project in an attempt to make the format as simple and effective as possible. There was a period of continuous and simultaneous testing and revising. Some of the decisions and problems encountered and procedures used during the creation and implementation were the determination of the specific structure details, the handling of source data, output arrangement, conversion of the original record, counting field lengths, and the use of the terminal and computer.

When the format components and the overall structure were finalized, a computer program in ALGOL language was written and keyed into the computer (see Appendix E). The program relates specifically to the MARGE format. A file of the sample Family Group Record data in MARGE format was created and stored in the Brigham Young University DECsystem-10 computer (see Appendix C). After the file and program were entered into the computer and debugged, the program was executed resulting in the printout of the genealogical data as shown in Figure 1.

```
HUSBAND: LAMB,DELBERT M. 03 03 1905 HYDE PARK, CCH, UT LAMB, GEORGE ZIMMERMAN&GRANT, JANE ELIZABETH 24 08 1932 SALT LA
KE CITY, S-LK, UT 04 03 1913 24 08 1932 24 08 1932 SL TEMPLE *NH*BD*BP*PH CH,BIRTH, SALT LAKE CITY, S-LK, UT/LDS ARCHIVES *M
D*MP VI,MARRIAGE, DELBERT M. LAMB, BOUNTIFUL, DAV, UT *TD CH,BAPTISM, HYDE PARK WARD, CCH, UT *ED*SD*SP TE,SL~

WIFE : WRIGHT, ALTA 04 06 1908 WOODS CROSS, DAV, UT WRIGHT, FRANCIS HENRY&STAHLE, SUSIE FANNY 04 06 1916 24 08 1932 24
08 1932 *NH*BD*BP VI,BIRTH, SALT LAKE CITY, S-LK, UT/HEALTH DEPT. *PW FA,BIBLE, ALTA W. LAMB, BOUNTIFUL, DAV, UT *TD CH,BAPTISM, WE
ST BOUNTIFUL WARD, DAV, UT *ED*SD TE,SL~

CHILD #1: LAMB, RONALD WRIGHT M 19 08 1934 SALT LAKE CITY, S-LK, UT 29 12 1958 BYWATER, IDA JOSEPHINE 05 07 1942 09
06 1954 BIC *NC*SC*BD*BP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT *MD*NS VI,MARRIAGE, RONALD W. LAMB, BOUNTIFUL, DAV, UT*TD CH,B
APTISM, BOUNTIFUL THIRD WARD, DAV, UT *ED*SD TE,SL~

CHILD #2: LAMB, RENEE F 03 05 1936 SALT LAKE CITY, S-LK, UT 27 06 1957 QUIGLEY, GORDON MYRL 05 08 1934 25 06 1957
BIC *NC*SC*BD*BP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT *MD*NS VI,MARRIAGE, GORDON M. QUIGLEY, SALT LAKE CITY, S-LK, UT*TD CH,BA
PTISM, BOUNTIFUL THIRD WARD, DAV, UT *ED*SD TE,SL~

CHILD #3: LAMB, LARUE F 08 12 1939 SALT LAKE CITY, S-LK, UT 18 08 1961 HAMBY, LOUIS AUSTIN 03 01 1938 17 05 1951 B
IC *NC*SC*BD*BP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT *MD*NS VI,MARRIAGE, L. AUSTIN HAMBY, CROSSVILLE, CUMBERLAND, TN*TD CH,BAP
TISM, BOUNTIFUL THIRD WARD, DAV, UT *ED*SD TE,SL~

CHILD #4: LAMB, CAROLYN F 26 07 1942 SALT LAKE CITY, S-LK, UT 02 06 1961 BEAN, GILBERT, MACK 05 08 1950 01 06 1961
BIC *NC*SC*BD*BP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT *MD*NS VI,MARRIAGE, GILBERT M. BEAN, BOUNTIFUL, DAV, UT*TD CH,BAPTISM,
BOUNTIFUL THIRD WARD, DAV, UT *ED*SD TE,SL~

CHILD #5: LAMB, MARILYN F 26 07 1942 SALT LAKE CITY, S-LK, UT 30 07 1942 BIC *NC*SC*BD*BP VI, BIRTH, ALTA W. LAMB, BOUNT
IFUL, DAV, UT *DD VI,DEATH, SALT LAKE CITY, S-LK, UT/HEALTH DEPT. *SD TE,SL~

CHILD #6: LAMB, RICHARD GRANT M 02 06 1944 SALT LAKE CITY, S-LK, UT 22 12 1975 MARTIN, NANCY CAROL 05 07 1952 24 01
1964 BIC *NC*SC*BD*BP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT *MD*NS VI,MARRIAGE, RICHARD G. LAMB, HOUSTON, HARRIS, TX*TD CH,BA
PTISM, BOUNTIFUL THIRD WARD, DAV, UT *ED*SD TE,SL~

CHILD #7: LAMB, CONNIE F 18 09 1947 SALT LAKE CITY, S-LK, UT 05 11 1955 09 10 1973 BIC *NC*SC*BD*BP VI,BIRTH, ALTA W,
LAMB, BOUNTIFUL, DAV, UT *TD CH,BAPTISM, BOUNTIFUL NINTH WARD, DAV, UT *ED*SD TE,SL~

SUBMISSION DATA: 16 11 1976 CONNIE LAMB 265 E. CENTER BOUNTIFUL, DAV, UT BOUNTIFUL, UTAH STAKE DAU DAU~
```

Figure 1

Printout Resulting from Computer Execution
on a MARGEDOS Record

On the printout, each field is identified with its field name and the data and sources follow the heading in paragraph form. The subfield codes are eliminated leaving spaces between data elements, and each field is set off from the others to facilitate reading the record. However the source codes remain so that what follows can be recognized as source information and the letters refer to the data elements in the order they appear in the field. If one source is used for more than one data element, all the codes for those elements are printed prior to the source. The dates are not included in the sources because they can be ascertained from the data. The sample printout is one possible output arrangement and could be altered by varying the program commands.

The total MARGEDOS record (machine-readable Family Group Record) is 2,829 characters long which amounts to a little over twice the number of characters on the original form. So in this case, providing sources for each piece of information and organizing the data with codes and tags required approximately double the space based on character count.

The author personally checked the sources for the data recorded on the sample family group form. Where possible, original records or copies of such were viewed or persons to whom the events occurred were contacted. High priority sources were used to assure accuracy of the record tested in this project. Each data element was documented with a source reference. The record utilized for this project is simple and most of the persons are still living so many of the sources were duplicated throughout the record. Although the sample record was

quite brief and straightforward, it served adequately as a tool to demonstrate the application of MARGEDOS.

The MARGEDOS format does not guarantee the accuracy of the data. Neither man nor machine can determine if the submitter has in fact checked the sources and supplied correct data. This project deals with authenticity, that is, the existence of the source documents. Providing source references gives the researcher an audit trail so that the sources may be evaluated. This does not diminish the need for using priority sources but does increase the need for careful recording of data and sources. The documentation of each data element is suggested as a means of establishing authenticity in order to achieve validity which is the internal accuracy of the genealogical record.

Implementation of MARGEDOS format on a large scale would require further study and adjustment. The current Family Group Record would need to be expanded to allow space for documentation of each name, date and place. Modification of the format may also be necessary to include aspects not considered in this project. There are many variations and possibilities that can occur in genealogical record-keeping that would need to be considered and problems may be encountered that were not foreseen in developing MARGEDOS. However, the format is flexible and alterations can easily be made. The editing process which includes coding, abbreviating, correcting, counting, and generally converting records into machine-readable form would need to be done by trained personnel. Provision would also be needed for updating the whole file periodically to correct mistakes or

include information that was originally missing and later supplied. For widespread application, the computer program should be general and have the capability to recall any length of record. So the current program would need to be revised or a new program written.

Another aspect of extended application would be the need for an educational program to teach patrons about the format and about aspects of documentation, such as how to compare sources, use of priority sources, and how to explicitly state a source for each data element. The submitter of the data would be expected to provide accurate information, as complete and detailed as necessary to be easily understood by persons editing the records.

MARGEDOS format is one possible way of using automation to manage the vast amount of genealogical data in existence and being gathered. Computers are most useful and effective in situations involving repetition, volume, urgency, and accuracy, all of which relate to genealogical data. Whatever form is used, control of genealogical records is essential. A MARC-like format allows for the storage and control of such records. They can be rapidly searched and any portion extracted or the entire record recalled from the computer. The meaning of the record and its reliability are most important. In MARGEDOS, the record data is clearly identified and its organization is well defined permitting ready communication of the meaning; the extensive documentation lends credibility to the record.

Chapter 6

CONCLUSION

The research project reported in this paper is of a creative nature rather than historical or statistical. It involved the production of a machine-readable format for genealogy based on the MARC bibliographic structure. The Family Group Record was employed as the basic record and its component elements were logically tagged and coded for computer manipulation. Special sections of the structure necessary for computer control of the data were developed from the record in machine-readable form. An integral aspect of the project was the inclusion of source documentation for each data element to improve reliability of the data contained within the genealogical records.

The formulation of the machine-readable format structure, called MARGEDOS, was followed by its application to a sample Family Group Record complete with data and sources. The record was converted into the MARGEDOS format and a computer program was written to recall the data. Execution of the program on the MARGEDOS file resulted in a printout of each person on the family group form with his or her genealogical data and the verifying sources printed together in sections.

For extensive implementation, modifications in the format, program, and output arrangement may be necessary or desired. The flexibility of the MARGEDOS format allows for alteration and

adaptation to meet various situations and needs. Application of the format on a large scale would involve a sizable amount of computer space and a definite financial investment. These aspects were not considered in this project and further study would need to be done on space and cost analysis and on modifications necessary for extensive application.

The feasibility of MARGEDOS format application was demonstrated by the computer execution on a complete and documented genealogical form. The concept of documentation is important to the authenticity and ultimately to the accuracy of the records. It is therefore proposed that the concepts and results of this project can be an aid to the control and utilization of genealogical records.

APPENDIX A

FAMILY GROUP RECORD FORM

FAMILY GROUP RECORD

PLACE S. Sharon, Windsor, Vt

ENTER ALL DATA IN THIS ORDER

Husband

Wife

Ward Examiners 1. 2.

Stake or Mission

NAME & ADDRESS OF PERSON SUBMITTING SHEET

RELATION OF ABOVE TO HUSBAND

RELATION OF ABOVE TO WIFE

FOUR GENERATION SHEETS FOR FILING ONLY

YES ☐ NO ☐

DATE SUBMITTED TO GENEALOGICAL SOCIETY

HUSBAND

Born Place

Chr. Place

Marr. Place

Died Place

Bur. Place

HUSBAND'S FATHER

HUSBAND'S MOTHER

HUSBAND'S OTHER WIVES

WIFE

Born Place

Chr. Place

Died Place

Bur. Place

WIFE'S FATHER

WIFE'S MOTHER

WIFE'S OTHER HUSBANDS

LDS ORDINANCE DATA

	BAPTIZED (Date)	ENDOWED (Date)	SEALED TO PARENTS (Date and Temple)
HUSBAND			
WIFE			

Sex M F	CHILDREN Given Names	SURNAME	WHEN BORN DAY	MONTH	YEAR	WHERE BORN TOWN	COUNTY	STATE OR COUNTRY	DATE OF FIRST MARRIAGE DAY MONTH YEAR TO WHOM	WHEN DIED DAY MONTH YEAR	BAPTIZED (Date)	ENDOWED (Date)	SEALED TO PARENTS (Date and Temple)
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													

SOURCES OF INFORMATION

OTHER MARRIAGES

NECESSARY EXPLANATIONS

©1972 The Genealogical Society of The Church of Jesus Christ of Latter-day Saints, Inc.

38

APPENDIX B

MARC-LIKE FORMAT STRUCTURE

APPENDIX B

MARC-LIKE FORMAT STRUCTURE

More detail about the MARC structure is presented here so that the reader may more easily understand the purpose, method, and results of this project. Perhaps the basic difference between a printed catalog record and one in machine-readable form is that much of the information on the former is expressed implicitly while the latter is often explicit. A human being can look at a printed card or record and recognize names, dates, places, and other data, as well as where each of these elements begins and ends. However, a machine cannot make such judgements and so a means must be provided by which it can discriminate between data elements and determine where each begins and ends. Once this has been done, the bibliographic data can be stored on a medium such as magnetic tape, and the computer can be programmed to manipulate the data elements in a variety of ways. This identification is necessary because the data is entered into the computer as a string of characters. A machine-readable system like MARC has a format in which the data elements are arranged precisely and labelled consistently much as on a catalog card so that the computer can identify the information.[1]

---

[1]U.S., Library of Congress, Information Systems Office, MARC Manuals used by the Library of Congress, Vol. 2, Data Preparation Manual: MARC Editors (3d ed.; Chicago: Information Science and Automation Division, American Library Association, 1970), pp. 1-3.

Definitions of Terms

Elements of records and components of the MARC format are defined as follows.[2]

Data Element. A unit of information.

Field. One or a group of data elements that are separate from each other but related. Each field is assigned a name which represents the contents of that field. There are two types of fields, fixed and variable. Each field is preceded by a field tag and followed by a terminator.

Record. A collection of fields treated as a unit.

Tag. A string of characters used to identify or name a field. Each is unique and consists of three numerals within the range of 001 to 999. The tag designates main content, but a field need not contain all possible data elements in every case.

Indicator. A one-character code which supplies additional information about the data field or parameters for the processing of the field. There may be more than one indicator per field, and they appear as the first items in each field.

Subfield Code or Data Element Identifier. A code consisting of one or more characters used to identify individual data elements

---

[2]Henriette D. Avram and Lucia J. Rather, "Principles of Format Design," Journal of Library Automation, 7 (September, 1974), 161. see also U.S., Library of Congress, MARC Manuals, op. cit., pp. 3-6, and Poon, op. cit., p. 15.

within a field. The code precedes the data element which it identifies. Usually lower case alphabetic characters are used in connection with a special symbol (delimiter) to designate the elements.

Delimiter. A special symbol such as a dollar sign used in conjunction with a letter to separate data elements in a field.

Terminator. A symbol used to indicate to the machine that it has reached the end of the data contained in that particular field.

Description of the Format Structure

The basic structure of the MARC format consists of a leader, a record directory, control fields, and variable fields. (See Figures 2-6.) A description of each is given as follows.[3]

Leader. A block of data that provides certain information about the ensuing record such as type of record and total length as well as the basic address of the data in the computer. It describes in brief the general make-up of the record.

Record Directory. A series of fixed-length entries which contain the identification tag, the length, and the starting character position of each variable field in the string of record data. With this information, any field can be extracted from the record as needed.

---

[3]Poon, op. cit., pp. 16-17. see also U.S., Library of Congress, Information on the MARC System, op. cit., pp. 2-5.

Control or Fixed Fields. Fields containing data elements which are fixed in length. They are always expressed by the same number of characters and are used to control record identification.

Variable Fields. Fields containing data that vary in length from record to record. This means that subfield codes and indicators are generally needed to denote the different data elements in each variable field. Each field begins with indicators specified for the type or function of the field. If indicators are not used with a particular field, the spaces will contain blanks. A subfield code, consisting of a delimiter symbol followed by a lower case letter, precedes each data element in the field and identifies the element.

SUMMARY OF THE MARC II COMMUNICATIONS FORMAT STRUCTURE

Leader	Record Directory	Control Fields	Variable Fields

A. LEADER -- The leader is fixed in length for all records and contains 24 characters.

B. RECORD DIRECTORY -- The record directory is made up of a series of fixed-length entries (12 characters each) which contain the identification tag, the length, and the starting character position in the record of each of the variable fields. The record directory will end with a field terminator code ($1E_{16}$, 8-bit; 36_8, 6-bit).

C. CONTROL FIELDS -- The control fields contain alphameric data elements, many of which have a fixed length. These fields end with a field terminator code. Each control field is identified by a 3-character numeric tag in the record directory, and these tags will not be repeated in a logical record.

D. VARIABLE FIELDS -- The variable fields are made up of variable length alphameric data, and all fields end with a field terminator code except the last variable field in a logical record where the field terminator is replaced by an end-of-record code ($1D_{16}$, 8-bit; 35_8, 6-bit). Each variable field is identified by a 3-character numeric tag in the record directory, and tags may be repeated as required in a logical record.

Figure 2

MARC Format Structure

Outline of Leader

Ø 4	5	6	7	8 9	10	11	12 16	17	20	23
Record Length	Status	Type of Record	Biblio. Level	Blanks	Indicator Count	Subfield Code Count	Base Address of Data	Encoding Level	Blanks	Entry Map

Element Number	Name of Leader Data Element	Number of Characters	Character Position in Record
1.	Logical Record Length	5	Ø-4
2.	Record Status	1	5
3.	Legend		
	a. Type of Record	1	6
	b. Bibliographic Level	1	7
	c. Blanks	2	8-9
4.	Indicator Count	1	10
5.	Sub-field Code Count	1	11
6.	Base Address of Data	5	12-16
7.	Encoding Level	1	17
8.	Blanks	2	18-19
9.	Entry Map (Proposed)	4	20-23

Contents of Leader

1. Logical Record Length -- the total number of characters in the logical record including itself. The number is right justified with leading zeros.

2. Record Status -- [n, c, or d]

   n = new record
   c = corrected or revised record
   d = deleted record

3. Legend

   a. Type of Record -- [a]

      a = language material, printed

   b. Bibliographic Level -- [m]

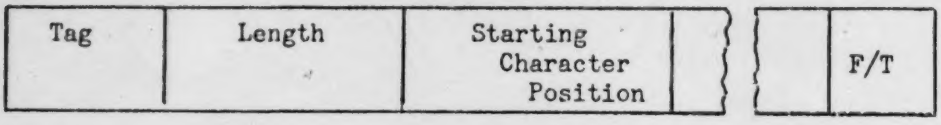
      m = monograph

   c. 2 blank characters

Figure 3

Leader

Outline of Record Directory Entries



F/T = Field Terminator

Element Number	Name of Record Directory Data Element	Number of Characters	Character Positions in Directory Entries
1.	Tag	3	Ø-2
2.	Field Length	4	3-6
3.	Starting Character Position	5	7-11

Contents of Record Directory Entries

1. Tag -- 3 numeric characters which identify the field. Tags in MARC II range from ØØ1 to 899.

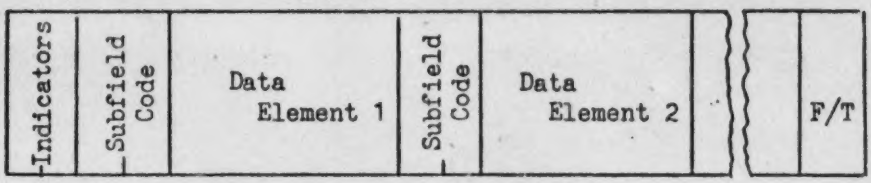
2. Field Length -- the number of characters in the field identified by the tag. This count includes indicators, subfield codes, data, and field terminator. The number is right justified with leading zeros.

3. Starting Character Position -- 5 numeric characters which give the character position in the record of the first character of the field. The character position is relative to a base which begins with the first character of the first field (i.e., for monographs the first character of the control number field). The first record directory entry will contain the starting character position ØØØØØ. Subsequent record directory entries will have starting character positions incremented by the field length of the previous entry.

Figure 4

Record Directory

Outline of Variable Fields



1. Indicators

   Each variable field will begin with 2 characters which provide descriptive information about the field. The contents of the indicators are specified for the fields in which they are used (see page 43ff). If the indicators are not used with a particular field, they will contain blanks.

2. Subfield Codes

   Variable fields are made up of a single data element or a group of data elements. The subfield code precedes each data element in a field and identifies the data element. The subfield code consists of 2 characters: a delimiter ($1F_{16}$, 8-bit; 37_8, 6-bit) followed by a lower case alphabetic character. (The 6-bit delimiter is in non-standard character set I and has a shift character 73_8 in front of it.) For the purpose of these specifications, the delimiter will be represented by the typewriter graphic "$." Subfield codes are specified for each variable field beginning on page 43.

3. Data Elements

   All data elements in the variable fields may have variable lengths.

4. Field Terminator

   The field terminator code is $1E_{16}$, 8-bit; 36_8, 6-bit. (The 6-bit field terminator is in non-standard character set I and has the shift character 73_8 in front of it.)

Figure 5

Variable Fields

SAMPLE LIBRARY OF CONGRESS CARD IN THE MARC II FORMAT

Leader | Record Directory

00441	n	a	m	ßß	2	2	00133	ßßßßßßß	001001300000	008004100013	050001600054

082001300070	100003500083	245008900118	260004800207	300001900255	650003400274	F

LC Card Number | Fixed Length Data Elements

ßßß	64	008443	ß	F	681001	s	1964	ßßßß	nyu	ßßßß	ß	ß	ßßßß	ß00100ß	eng	0	ß	F

LC Call Number | Dewey Number | Main Entry

0ß	$aPS614$b.L75	F	ßß	$a811.5082	F	1Ø	$aLowenfe s,ßWalter,$d1897-$eed.	F

Title

1ß	$aPoetsßofßtoday;$baßnewßAmericanßanthology.$cWithßaßprologueßpoemßbyß

Imprint

Langstonß Hughes.	F	0ß	$aNewßYork,$bInternationalßPublishers$c [1964]	F

Collation | Subject Heading

ßß	$al43ßp.$c21ßcm.	F	ß0	$aAmericanßpoetry$y20thßcentury	R

ß = blank F = field terminator R = end of record

Lowenfels, Walter, 1897– *ed.*

Poets of today; a new American anthology. With a prologue poem by Langston Hughes. New York, International Publishers [1964]

143 p. 21 cm.

1. American poetry—20th cent. I. Title.

PS614.L75 811.5082 64–8443

Library of Congress ◯ [5]

Figure 6

Sample MARC Record

APPENDIX C

SAMPLE MARGEDOS RECORD

P2529 3P0142 1960364P0P0P020P0298P03A65P1P3P6P0A64P0P03P5P0497P5P3P3P0P1275500P0290015795P5P1990187750603P5W2076527P207P2381P02
P2497258P" 3WMLAMH,DELBERT M,36D03 03 1935SHPHYDE PARK, CCH, UT$PHLAMH, GEORGE ZIMMERMAN&GRANT, JANE ELIZABETHSP024 08 1932SHPSALT
LAKE CITY, S-LK, UT$TD04 03 1913SED24 08 1932$3D24 08 1932$SPSL TEMPLE #NH#PD#HP#PH CH,BIRTH, SALT LAKE CITY, S-LK, UT/LDS ARCHIVES
#PD#HP VI,MARRIAGE, DELBERT M. LAMB, BOUNTIFUL, DAV, UT #TD CH,BAPTISM, HYDE PARK WARD, CCH, UT #ED#SD#SP TE,SL"$WWRIGHT, ALTA$D04
26 1928$HP#HOOS CROSS, DAV, UT$P#WRIGHT, FRANCIS HENRY&STAHLE, SUSIE FANNY$TD04 06 1916$ED24 08 1932$3D24 08 1932 #N#8D#HP VI,BIRT
H, SALT LAKE CITY, S-LK, UT/HEALTH DEPT. #DH FA,BIBLE, ALTA P. LAMB, BOUNTIFUL, DAV, UT #TD CH,BAPTISM, WEST BOUNTIFUL WARD, DAV, UT
#LD#SD TE,SL"3NCLAMH, RONALD WRIGHTSCFHD19 04 1934$HPSALT LAKE CITY, S-LK, UT$HD29 12 1958$NSHYRATER, IDA JOSEPHINE$TD05 07 19425
ED#9 26 1958$3D$IC #NC#SC#HD#HP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT #M0#NS VI,MARRIAGE, RONALD W. LAMB, BOUNTIFUL, DAV, UT#TD
CH,BAPTISM, BOUNTIFUL THIRD WARD, DAV, UT #ED#SD TE,SL"3NCLAMH, RENEESCFHD03 06 1936$HPSALT LAKE CITY, S-LK, UT$HD27 06 1957$NSGU
IGLEY, GORDON MYRL$TD05 08 1940$HD25 08 1957$3D$IC #NC#SC#HD#HP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT #M0#NS VI,MARRIAGE, GORDO
N M, GUIGLEY, SALT LAKE CITY, S-LK, UT#TD CH,BAPTISM, BOUNTIFUL THIRD WARD, DAV, UT #ED#SD TE,SL"3NCLAMB, LARUESCFHD08 12 1939$5PS
ALT LAKE CITY, S-LK, UT$HD18 08 1961$NSHAMBY, LOUIS AUSTIN$TD03 01 1940$ED17 08 1961$3D$IC #NC#SC#HD#HP VI,BIRTH, ALTA W. LAMB, BOUN
TIFUL, DAV, UT #M0#NS VI,MARRIAGE, L. AUSTIN HAMBY, CROSSVILLE, CUMBERLAND, TN#TD CH,BAPTISM, BOUNTIFUL THIRD WARD, DAV, UT #ED#SD T
E,SL"3NCLAMH, CAROLYNSCFHD26 07 1942$5FSALT LAKE CITY, S-LK, UT$HD02 06 1961$NSHEAN, GILBERT, MACK$TD05 08 1950$ED01 08 1961$3D$IC
#NC#SC#3D#HP VI,BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT #M0#NS VI,MARRIAGE, GILBERT M. HEAN, BOUNTIFUL, DAV, UT#TD CH,BAPTISM, BOUN
TIFUL THIRD WARD, DAV, UT #ED#SD TE,SL"3NCLAMB, MARILYNSCFHD26 07 1942$HPSALT LAKE CITY, S-LK, UT$HD30 07 1942$3D$IC #NC#SC#3D#P
VI, BIRTH, ALTA W. LAMB, BOUNTIFUL, DAV, UT #DD VI,DEATH, SALT LAKE CITY, S-LK, UT/HEALTH DEPT, #SD TE,SL"3NCLAMH, RICHARD GRANT$SCH
$3D22 06 1960$HPSALT LAKE CITY, S-LK, UT$HD22 12 1975$NSMARTIN, NANCY CAROL$TD05 07 1952$ED24 01 1964$3D$IC #NC#SC#3D#HP VI,BIRTH, A
LTA W. LAMB, BOUNTIFUL, DAV, UT #M0#NS VI,MARRIAGE, RICHARD G. LAMB, HOUSTON, HARRIS, TX#TD CH,BAPTISM, BOUNTIFUL THIRD WARD, DAV, U
T #ED#SD TE,SL"3NCLAMB, CONNIESCFHD18 09 1947$HPSALT LAKE CITY, S-LK, UT$TD05 11 1955$ED09 12 1973$3D$IC #NC#SC#3D#HP VI,BIRTH, AL
TA W. LAMB, BOUNTIFUL, DAV, UT #TD CH,BAPTISM, BOUNTIFUL NINTH WARD, DAV, UT #ED#SD TE,SL"3DS16 11 1976$NSCONNIE LAMB$AD265 E, CENTE
R BOUNTIFUL, DAV, UT$WMBOUNTIFUL, UTAH STAKE$RHDAUGRWDAU"

APPENDIX D

TYPES OF SOURCES

Table 3

Types of Sources and Specifications*

The following lists give representative examples of types of records within each source class, their possible origins (taken from), and the specifications for use. The class and types should be used by the submitter when recording sources. The origin should usually be given as part of the location.

Class and types	Taken from	Specifications
	VITAL	
Birth Marriage Death	Certificates of government (federal, state, county, town, or municipal)	Name Birthdate Birthplace Parents' names Marriage date Marriage place Death date Death place
	CHURCH	
Christening Baptism Marriage Burial	Certificates, accounts, membership rolls, etc. of parish registers, ward records (LDS), patriarchal blessings, and other religious records	Use to verify names, dates, and places of events
	CIVIL	
Legal (court, other than wills)	Records and transactions of same	May use to document any data

Table 3 (continued)

Class and types	Taken from	Specifications
Land Archives Tax records	Records and transactions of same	May use to document any data
	FAMILY	
Bible Diaries, journals, letters Photo albums, scrapbooks Household items Personal communication	Same as types listed	May use to document any data (must be specific relationship to use)
	MISCELLANEOUS	
A. Single Military Sexton Mortician, coroner Tombstones & printed cemetary Obituaries Passport, immigration, passenger lists, naturalization Associations (family, state)	Same as types listed	May use to document any data

Table 3 (continued)

Class and types	Taken from	Specifications
Directories (city, telephone, trade, professional) Insurance Deeds, mortgages	Same as types listed	May use to document any data
B. Composite Printed family histories & genealogies Printed place histories		
CENSUS		
Federal	Federal census record of U.S. - 1850 & 1860 and Canada - 1851 & 1861	See Submission Manual for restriction of use Use for childrens' data only Direct-line only
PROBATE		
Wills Affidavits, etc.	Court records	See Submission Manual for restrictions
PERSONAL KNOWLEDGE		
For the purposes of this project, this means "hearsay" - no definite record is used		May be used for any data

Table 3 (continued)

Class and types	Taken from	Specifications
	TEMPLE	
Baptism Endowment Sealing	Temple records of ordinance performance	Used for temple work only

*Based on Records Submission Manual, op. cit., pp. 3-5, and Stevenson, Noel C., Search and Research, the Researcher's Handbook (rev. ed., Salt Lake City: Deseret Book, 1959).

APPENDIX E

COMPUTER PROGRAM

```
BEGIN
STRING MARGEDATA, DOLLAR, BLANK, SS, AT;
STRING ARRAY HDR[1:10];
INTEGER I,S,RECLENGTH,J,K,BASE,CODE;
INTEGER ARRAY LENGTH,FLAG,POSITION,MARGE[1:10];
FOR I←1 UNTIL 10 DO
    BEGIN
    LENGTH[I]←0; POSITION[I]←0; FLAG[I]←0;
    END;
    RECLENGTH←0; BASE←0; CODE←0; DOLLAR←"$"; BLANK←" "; AT←"@";
MARGE[1]←100; MARGE[2]←288; MARGE[3]←521; MARGE[4]←592;
MARGE[5]←593; MARGE[6]←594; MARGE[7]←595; MARGE[8]←596;
MARGE[9]←597; MARGE[10]←800;
HDR[1]←"HUSBAND: "; HDR[2]←"WIFE: "; HDR[3]←"CHILD #1: ";
HDR[4]←"CHILD #2: "; HDR[5]←"CHILD #3: "; HDR[6]←"CHILD #4: ";
HDR[7]←"CHILD #5: "; HDR[8]←"CHILD #6: "; HDR[9]←"CHILD #7: ";
HDR[10]←"SUBMISSION DATA: ";
WRITE("PLEASE TYPE YOUR FILE NAME--ENCLOSED IN QUOTES"); NEWLINE;
READ(SS);
INPUT(5,"DSK");
SELECTINPUT(5);
OPENFILE(5,SS);
FOR I←1 UNTIL 5 DO
    BEGIN
    INSYMBOL(S);
    RECLENGTH←RECLENGTH + (S - 48)*10^(5-I);
    END;
WRITE("RECLENGTH="); PRINT(RECLENGTH); NEWLINE;
MARGEDATA←NEWSTRING(RECLENGTH,7);
FOR I←6 UNTIL RECLENGTH DO
    BEGIN
    INSYMBOL(S);
    MARGEDATA.[I]←S;
    END;
WRITE("THE MARGE RECORD HAS BEEN READ FROM DISK"); NEWLINE(2);
FOR I←12 UNTIL 16 DO BASE←BASE+(MARGEDATA.[I]-48)*10^(16-I);
WRITE("BASE="); PRINT(BASE); NEWLINE(3);
FOR I←22 STEP 12 UNTIL BASE + 1 DO
    BEGIN
    FOR J←I UNTIL I+2 DO CODE←CODE+(MARGEDATA.[J]-48)*10^(I+2-J);
    FOR K←1 UNTIL 10 DO IF CODE=MARGE[K] THEN GO TO MATCH;
    WRITE("NO MATCH ON "); PRINT (CODE); NEWLINE; GO TO NOMATCH;
    MATCH:FOR J←I+3 UNTIL I+6 DO
        LENGTH[K]←LENGTH[K]+(MARGEDATA.[J]-48)*10^(I+6-J);
    FOR J←I+7 UNTIL I+11 DO
        POSITION[K]←POSITION[K]+(MARGEDATA.[J]-48)*10^(I+11-J);
    FLAG[K]←1;
    NOMATCH:  CODE←0
    END;
FOR K←1 UNTIL 10 DO IF FLAG[K]=1 THEN
    BEGIN
    WRITE(HDR[K]); HDR[K]←NEWSTRING(LENGTH[K],7);
    S←LENGTH[K];
    LENGTH[K]←LENGTH[K] + BASE + POSITION[K];
    HDR[K].COPY(MARGEDATA,BASE+POSITION[K]+2,LENGTH[K]);
    FOR J←1 UNTIL S-4 DO
        BEGIN
        IF HDR[K].[J]=DOLLAR.[1] THEN GO TO LOOP;
        HDR[K].[J]←BLANK.[1];
        HDR[K].[J+1]←BLANK.[1];
        HDR[K].[J+2]←BLANK.[1];
        LOOP: END;
    WRITE(HDR[K]);  NEWLINE(2);
    END;
PRINT(RECLENGTH); WRITE(MARGEDATA);
RELEASE(5);
END
```

SOURCES CONSULTED

SOURCES CONSULTED

Books

American Society of Genealogists. Genealogical Research Methods and Sources. Vol. 1, Milton Rubincam, ed. Washington, D.C.; n.p., 1960.

Church of Jesus Christ of Latter-day Saints. Genealogical Society. Records Submission Manual. 4th ed. Salt Lake City: Publishers Press, 1973.

Church of Jesus Christ of Latter-day Saints. Genealogical Society. Ten Basic Steps in Genealogy. Salt Lake City: Deseret Book, 1965.

Church of Jesus Christ of Latter-day Saints. Priesthood Genealogy Committee. A Continuing Priesthood Program for Family Exaltation. [Salt Lake City] 1970.

Dine, L. G. The Genealogist's Encyclopedia. New York: Weybright and Talley, 1969.

Fudge, George H. and Frank Smith. LDS Genealogist's Handbook; Modern Procedures and Systems. Salt Lake City: Bookcraft, 1972.

Harland, Derek. Genealogical Research Standards. Salt Lake City: Bookcraft, 1953.

Irvine, Ruth. MARC for Cataloguers: an Explanation of Its Use. Southampton University of Southampton, 1972.

Jaussi, Laureen and Gloria D. Chaston. Fundamentals of Genealogical Research. Salt Lake City: Deseret Book, 1972.

Jones, Vincent L., Arlene H. Eakle and Mildred H. Christensen. Genealogical Research; a Jurisdictional Approach. Rev. ed. Salt Lake City: Publishers Press, 1972.

Stevenson, Noel C. Search and Research, the researcher's handbook; a guide to official records and library sources for investigators, historians, genealogists, lawyers and librarians. Rev. ed. Salt Lake City: Deseret Book, 1959.

United States. Library of Congress. Information Systems Office. MARC Manuals used by the Library of Congress. 2d ed. Chicago: Information Science and Automation Division, American Library Association, 1970.

United States. Library of Congress. MARC Development Office. Information on the MARC System. Washington, D.C.: Library of Congress, 1973.

World Conference on Records and Genealogical Seminar, Salt Lake City, 1969. Papers. Vol. 2. Salt Lake City: Genealogical Society of The Church of Jesus Christ of Latter-day Saints, 1969.

Periodicals

Avram, Henriette D. and Kay D. Guiles. "Content Designators for Machine-readable Records: a Working Paper," Journal of Library Automation, 5 (December, 1972), 207-216.

Avram, Henriette D. and Lucia J. Rather. "Principles of Format Design," Journal of Library Automation, 7 (September, 1974), 161-167.

Burton, Theodore M. "I Have a Question," Ensign, 5 (March, 1975), 32-33.

Goossens, Paula. "Techniques for Special Processing of Data within Bibliographic Text," Journal of Library Automation, 7 (September, 1974), 168-182.

Unpublished Sources

Poon, Hilda. "MARC-Saf, the MARC-like Invoicing Format for Subscribers and Publishing Agents." Unpublished research paper, Brigham Young University, 1976.

Stirling, Keith. Personal correspondence from Mr. Keith Stirling, Brigham Young University to Mr. Charles Clement, Genealogical Library, Salt Lake City. April 13, 1976.